Bulletin UASVM Agriculture, 67(1)/2010
Print ISSN 1843-5246; Electronic ISSN 1843-5386

# Pedometric Techniques in Spatialisation of Soil Properties for Agricultural Land Evaluation

**Iuliana Cornelia TANASĂ[1)2)], Mihai NICULITĂ[1)], Bogdan ROŞCA[3)], Radu PÎRNĂU[1)4)]**

[1)] Faculty of Geography and Geology, University Al. I. Cuza,
Carol I. Avenue, No. 20A, 700505, Iaşi, Romania
[2)] Data Invest Company, A Panu Street, 26, World Trade Center Iaşi, Romania;
[3)] Geography Group, Iaşi Branch, Romanian Academy, Carol I Avenue, Iaşi, Romania;
[4)] Office for Pedological and Agrochemical Studies, Iaşi, Romania;
e-mail: iuliananiculita@geomorphologyonline.com; iuliana.tanasa@datainvest.ro;
roscao@gmail.com; radupirnau@yahoo.com

**Abstract.** Agricultural land evaluation plays an important role in the pedologic and economic foundation of a sustainable agricultural practice and management. Usually, the agricultural land evaluation framework use soil, environment and land improvements data for computing an index of suitability for various agricultural uses.

Because of the punctual availability of soil data, we investigated the use of pedometric and geostatistical techniques (multiple linear regression, logistic regression, kriging) for their spatialisation, to be used further, for computing the suitability index, according to Romanian Agricultural Land Evaluation Methodology. The pedometric techniques were applied to a soil legacy database (620 described soil profiles with analytical data covering 15 villages from Iasi County, made by Iaşi Office for Pedology and Agrochemistry. The results are promising, but the quality of results, depend on the quality and quantity of soil data.

**Keywords**: pedometry, geostatistic, soil properties, agricultural land evaluation.

## INTRODUCTION

Agricultural land evaluation plays an important role in the pedologic and economic foundation of a sustainable agricultural practice and management (Smyth and Dumanski, 1993).

One of the most important aspect of land evaluation is the quality of soil survey data (Davidson, 2002). The majority of soil legacy data consist of punctual soil profile properties, and the classical land evaluation approach in romanian methodology rely on the soil profile to soil polygon generalization of soil data (Florea *et al*, 1987). The rise of pedometric techniques (Webster, 1994, McBratney *et al.*, 2000) give the possibility to spatialise the soil profile properties, obtaining continuous data, which can be extended to the land evaluation approach (McBratney *et. al.*, 2000), giving the possibility to compute the suitability index in an spatialised manner. The most suitable approach is the use of pedometrical geostatistic () and soil covariates. The most used and cited pedometric techniques, are kriging, linear/logistic regression and regression kriging (Webster, 1994, Odeh *et al*, 1995, McBratney *et al*, 2000, Hengl *et al*, 2004b, Hengl *et al*, 2007).

The Romanian methodology for land evaluation (Florea *et al*, 1987) take in consideration 18 variables, from which 2 are climatic, 2 hydrologic, 3 geomorphologic, 1 human related and 10 pedologic. For every variable as a function of its value, a coefficient,

varying from 0 to 1 is attributed after several tables, based on the membership of the variable value in a specified class.

The use of pedometric approach of soil property interpolation give the possibility to use the pixel as the homogenous mapping unit, for land evaluation. The main question is how much the legacy data supports a meaningful pedometric modeling. In this study we investigate the possibility of using the pedometric and geostatistic techniques and soil covariates for interpolating soil properties between the known data points, for the following variables: gleization, concerning the probability of presence of a gleyic horizon, salinization and alkalization, concerning the probability of presence of a salinic/natric horizon, soil texture in the first horizon, soil $CaCO_3$ in the first horizon, reaction in the first horizon, soil depth (used to obtain the soil volume), humus reserve in the first horizon. These approaches don't cover all the inputs for the Romanian methodology, but the pedometric methodology can be extended to all the required inputs.

## MATERIALS AND METHODS

The geostatistical analyses (kriging, regression kriging with logistic regression, crossvalidation) were applied in R (Ihaka and Gentleman, 1996), using base function of R and several packages: gstat and dependencies (Pebesma, 2004), mlogit and dependencies.

For the regression kriging (Odeh *et al*, 1995, Hengl *et al*, 2004b, Hengl *et al*, 2007) we used an R script (Hengl, 2009) and gstat. For all the geomorphometrical parameters entered in the regressions the significance levels smaller than 0.05 were chosen.

The presented pedometric methodology was applied to a soil legacy database (620 soil profiles) made by Romanian Office for Pedology and Agrochemistry, Iasi County, between 1989-2005. These data cover 15 villages from Iasi County (Fig. 1). The physical and chemical analysis of soil samples were made according to Romanian methodology (Florea, *et al*, 1987).

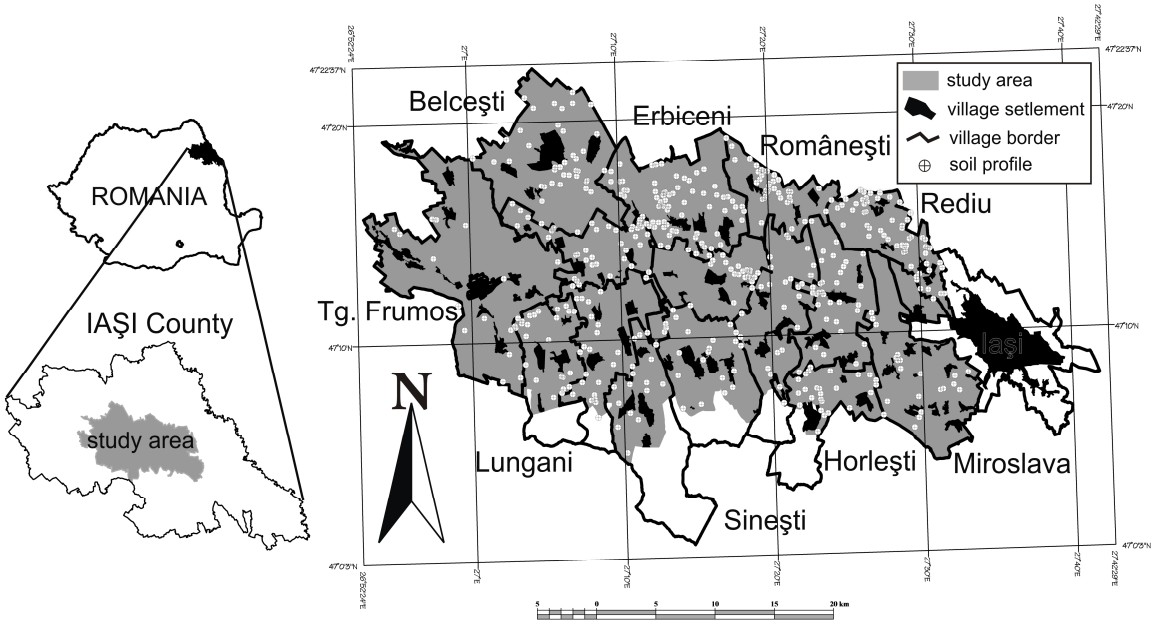


Fig. 1. Study area

The covariates for pedometric digital soil mapping were comprised of 29 geomorphometrical variables computed from a Digital Elevation Model. The DEM was interpolated from digitized contours from 25k topographic maps (1984-1985 edition) and height points from 5k maps (1964-1965 and 1979-1991 editions). As interpolators we used splines from GRASS (GRASS, 2010), multilevel splines from SAGA, minimum curvature from TntMips and Topo to Raster from ArcGIS, and we averaged the four DEMs, in order to reduce the "padi terraces" (Hengl *et. al.*, 2004a) and other artifacts. The pixel size was chosen to be 12.5 m, half the minimum representable (Hengl, 2006) width for 25k scale (1 mm for 25k means 25 m). The majority of geomorphometrical parameters were derived in SAGA with standard settings from the 12.5 m DEM. Only the maximum gradient was obtained using Tnt Mips 7.3 as the biggest slope between the center of a 3x3 pixels window and their 8 neighbour. The catena landform classification was obtained, using altitude above channel (<1 m) to separate the floodplains, then catchment area (< 1000) to separate the ridges.

The hillslopes class was assigned to the areas not being floodplains and ridges. Then, using Cluster Analysis for Grids module from SAGA, and the Hill-Climbing algorithm with 10 classes, normalization and the following rasters: maximum gradient, catchment area, MRVBF, the step hillslopes were separated. Using subtract vector combination in TntMips 7.3, we obtained the final step hillslopes without flooplains, and ridges without steep hillslopes. The areas not being floodplains, steep hillslopes and ridges, was assigned to gentle hillslopes, resulting 4 classes of catena landforms, according to the cuesta landform type of the area.

## RESULTS AND DISCUSSION

The regression parameters of the used multiple logistic regression modeling in the regression kriging are presented in Tab. 1. It can be seen that the strength of the correlation decrease with the decrease of cases of presence of the horizons. Although the adjusted $R^2$ value is not so high, the regression has significance. The other soil variables used in this study didn't had significant correlations with geomorphometrical covariates. This situation show that the legacy database is not very appropriate for error less geostatistical modeling.

Tab. 1

The parameters of the regressions for the considered soil properties

| Soil property | Equation | No. of cases/no. of presence cases for the logistic regression | $R^2_{adj.}$ |
|---|---|---|---|
| Probability of a gleyic horizon presence | $1/(1+e^{-z})$, where z = 0.03456 – 0.2765 * MG + 0.03052 * SH – 5.10261 * NH + 8.13453 * CA – 0.05055 * AACN + 0.53234 * CLC | 620/159 | 0.34 |
| Probability of a salinic horizon presence | $1/(1+e^{-z})$, where z = -1.62047 – 0.04031 * AACN + 0.06638 * WI | 620/137 | 0.19 |
| Probability of a natric horizon presence | $1/(1+e^{-z})$, where z = -3.217536 - 0.028002 * AACN + 0.101476 * WI | 619/50 | 0.09 |

Note: For geomorphometrical parameters we used the following codes, HOFD = Horizontal Overland Flow Distance, AACN = Altitude Above Channel Network, SH = Standardized height, NH = Normalized Height, CLC = Catena Landform Classification, CA = Catchment Area, MG = Maximum gradient, WI = Wetness Index, and the variables entered in the regressions had level of significance <0.05.

The parameters of the semivariogram modeling used in the ordinary kriging are presented in Tab. 2. It can be seen that the majority of the cases of kriging have significant nugget-sill ratios.

Tab. 2

Parameters of the semivariogram modeling for ordinary kriging (OK) and regression kriging (RK)

| | Variogram model | nugget | sill | range | Nugget -sill ratio (%) |
|---|---|---|---|---|---|
| Probability of a gleyic horizon presence (RK) | exponential | 0.82 | 0.90 | 288.61 | 91 |
| Probability of a salinic horizon presence (RK) | exponential | 0.85 | 0.94 | 2831.97 | 90 |
| Probability of a natric horizon presence (RK) | exponential | 0.20 | 0.52 | 1286.73 | 38 |
| Clay content (%) of the first horizon (OK) | exponential | 59.20 | 79.86 | 750.22 | 74 |
| Silt content (%) of the first horizon (OK) | exponential | 18.85 | 32.78 | 7680.92 | 57 |
| Sand content (%) of the first horizon (OK) | exponential | 65.78 | 101.08 | 1280.62 | 65 |
| $CaCO_3$ (%) in the first horizon (OK) | exponential | 4.52 | 7.33 | 1416.27 | 62 |
| Reaction in the first horizon (OK) | exponential | 0.16 | 0.67 | 684.59 | 24 |
| Soil depth (cm) (OK) | exponential | 422.61 | 728.14 | 7144.37 | 58 |
| Humus content (%)in the first horizon (OK) | exponential | 1.22 | 1.87 | 625.33 | 65 |

The cross-validation results of the ordinary and regression kriging are presented in Tab. 3. It can be seen that the cross-validation errors are quite high, altough the kriging modeling seems to be good. Their distribution seems to be normal (the mean and the median are very close one to each other). Probably the unequal spacing of the value points generate so high errors in the cross-validation approach.

Tab. 3

Cross-validation results for the used pedometric techniques (RK and OK)

| | Min | Max | Mean | Median |
|---|---|---|---|---|
| Probability of a gleyic horizon presence (RK) | 0 | 1 | 0.26 | 0.21 |
| Probability of a salinic horizon presence (RK) | 0 | 0.73 | 0.22 | 0.19 |
| Probability of a natric horizon presence (RK) | 0 | 0.79 | 0.08 | 0.03 |
| Clay content (%) of the first horizon (OK) | 30.31 | 49.22 | 38.51 | 38.51 |
| Silt content (%) of the first horizon (OK) | 17.4 | 31.27 | 24.40 | 24.50 |
| Sand content (%) of the first horizon (OK) | 22.16 | 49.63 | 36.75 | 36.76 |
| $CaCO_3$ (%) in the first horizon (OK) | 0.27 | 5.8 | 1.69 | 1.66 |
| Reaction in the first horizon (OK) | 5.33 | 8.73 | 7.28 | 7.28 |
| Soil depth (cm) (OK) | 50.55 | 141.25 | 89.42 | 91.13 |
| Humus content (%) in the first horizon (OK) | 2.13 | 5.59 | 3.66 | 3.65 |

## CONCLUSIONS

The spatialisation of the soil properties using the presented pedometric and geostatistic methods applied to the Romanian legacy soil database has promising and significant results, but the final error estimation and uncertainty depends strongly on the quality of soil data. The quality refers to the sampling schema which might not be optimal for kriging or for regression. There is need for an analysis of acceptable uncertainty of pedometric modeled soil data in land evaluation estimation.

*Acknowledgments.* This work was supported by the European Social Fund in Romania, under the responsibility of the managing Authority for the Sectoral Operational Programme for Human Resources Development 2007-2013 [grant POSDRU/88/1.5/S/47646 and grant POSDRU/89/1.5/S/49944].

## REFERENCES

1. Davidson, D. (2002). The assessment of land resources: Achievements and new Challenges. Australian Geographical Studies. 40:109-128.

2. Florea, N., V. Bălăceanu, C. Răută and A. Canarache (1987). Methodology of pedological studies elaboration. III volumes. Institute of Pedological and Agrochemical Research Press. Bucharest. (in romanian).

3. GRASS Development Team. (2010). Geographic Resources Analysis Support System (GRASS) Software. Version 6.4.0. Open Source Geospatial Foundation. http://grass.osgeo.org

4. Hengl, T. S. Gruber and D. P. Shrestha. (2004a). Reduction of errors in digital terrain parameters used in soil-landscape modeling. International Journal of Applied Earth Observation and Geoinformation. 5:97-112.

5. Hengl, T. G. B. Heuvelink and A. Stein. (2004b). A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma. 120:75-93.

6. Hengl, T. (2006). Finding the right pixel size. Computers & Geosciences. 32:1283-1298.

7. Hengl, T. G. B. Heuvelink and D. G. Rossiter. (2007). About regression-kriging: from equations to case studies. Computers & Geosciences. 33:1301-1315.

8. Hengl, T. (2009). A practical guide to geostatistical mapping. 2nd Edt. University of Amsterdam. www.lulu.com. p. 291.

9. Ihaka, R. and R. R. Gentleman. (1996). A language for data analysis and graphics. Journal of Computational and Graphical Statistics. 5:299-314.

10. McBratney, A., I. Odeh, T. Bishop, M. Dumbar and M. Shatar. (2000). An overview of pedometric techniques for use in soil survey. Geoderma. 97:293-327.

11. Odeh, I. O. A., A. B. McBratney and D. J. Chittleborough. (1995). Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma. 67: 215–226.

12. Pebesma, E. (2004). Multivariable geostatistics in S: the gstat package. Computers & Geosciences. 30:683-691.

13. Smith, A. J. and J. Dumanski. (1993). FESLM: an international framework for evaluating sustainable land management. World Soil Resources Report. FAO.

14. Webster, R. (1994). The development of pedometrics. Geoderma. 62:1-15.